Specifity determinants for protein secretion in Bacillus subtilis.

Jongbloed, Jan Douwe Harm

Chapter 10

Summary and general conclusions

In both prokaryotes and eukaryotes, a large number of proteins that are synthesised in cytoplasmic compartments, are transported across membranes in order to reach their final destination and fulfil their function. Because of the hydrophobic nature of membranes and the necessity to maintain their integrity, complex and strictly organised protein translocation machinery have evolved. Transported proteins are usually synthesised as precursors containing amino-terminal export signals. These signals contain information about the destination of these proteins and the transport machinery that should be used to reach their destination. During or immediately after synthesis of pre-proteins, soluble chaperones and/or targeting factors recognise such proteins and assist in targeting of the pre-protein to the cytoplasmic membrane. After arriving at the membrane and association with the correct translocation machinery, these proteins are transported through a proteinaceous channel. If the protein is translocated in an unfolded state, it folds into its final conformation after leaving the translocase. Depending on the nature of the export signal, the protein is then either retained in the membrane or released upon processing by specialised signal peptidases (SPases).

In the Gram-positive bacterium *Bacillus subtilis* at least four pathways for protein export are known. The largest number of secretory proteins appear to be transported *via* the general Sec (protein secretion) pathway. In contrast to the Sec machinery, which has a general function in protein transport, small numbers of proteins are exported *via* dedicated transport pathways. For example, proteins that are involved in the uptake of DNA during competence development use a pseudopilin export machinery, generally known as the Com pathway. In addition, some peptide antibiotics and pheromones are secreted *via* ATP-binding cassette (ABC) transporters. Finally, proteins can be exported *via* the twin-arginine translocation (Tat) pathway. This pathway has probably evolved to facilitate the translocation of proteins that have to incorporate a cofactor in the cytoplasm, or proteins that simply fold too rapidly or too tightly to be handled by the Sec machinery. Based on specific features of their respective signal peptides, the transport route of secretory pre-proteins can be predicted. While the signal peptides of proteins transported *via* the Com pathway or ABC transporters differ significantly from Sec-type or twin-arginine (RR-) signal peptides, the latter two share a characteristic tripartite structure. During the transport process *via* the Sec and/or Tat pathways, the recognition of the respective pre-proteins by specific pathway components is important for the subsequent steps in protein translocation. Thus, specificity determinants play key roles during protein export *via* these transport routes of *B. subtilis*.

First, the recognition by various chaperones and/or targeting factors is important. This determines whether the protein will fold, or will be kept in an unfolded conformation during its sojourn in the cytoplasm. The latter is required for transport *via* the Sec pathway. Second, the specific binding to chaperones and/or targeting factors, as well as the nature of the signal peptides of the transported proteins, determine pathway choice. The majority of proteins synthesised with Sec-type signal peptides will be targeted to the Sec machinery. However, a subset of proteins that have Sec-type signal peptides but, in addition, contain a so-called RR- or KR-motif have the potential to travel *via* the Tat pathway. Third, receptor-like components of the translocase select precursors for subsequent translocation. For example, the Ffh protein has an important signal peptide receptor function in the transport of proteins that are exported from the cytoplasm *via* the Sec pathway. Likewise, the signal peptides of proteins that travel *via* the Tat pathway may be recognised specifically by one

of the two paralogous TatC proteins of *B. subtilis*. Fourth, if paralogous copies of translocation machinery components are present, specific recognition of signal peptides and mature parts of the protein is important for effective transport of certain proteins. This does not apply to the Sec machinery of *B. subtilis*, since this organism does not contain paralogous copies of Sec translocase components. In contrast, the presence of paralogous TatA and TatC components that function in Tat-dependent translocation in *B. subtilis*, suggests that these components might play an important role as specificity determinant at the level of the translocase. Fifth, in organisms with multiple SPases, these specialised peptidases often show differences in specificity for subsets of precursors. The availability of the large number of five different SPases in *B. subtilis* suggests that at least some of the individual proteins have specialised functions. Sixth, the presence of various folding catalysts implies that their activities towards different secreted proteins vary with the nature of the mature protein. Finally, when post-translocational modifications of secreted proteins are needed, again components with different specificities may be involved (*i.e.* disulphide bond-forming proteins).

In this thesis, the role of specificity determinants for protein transport in *B. subtilis* is described. Especially, determinants that function at the level of pathway choice (signal peptides and translocase machinery components) and precursor processing (SPases) were studied and are discussed.

In Chapter 2, experiments are described to investigate whether the unique membrane anchor domain of *Bacillus* SPases is required for *in vitro* activity. For this purpose, soluble forms of SipS of *B. subtilis*, SipS of *Bacillus amyloliquefaciens* and SipC of the thermophile *Bacillus caldolyticus*, which lack their unique amino-terminal membrane anchor domain, were constructed. Of these three proteins, only a hexa-histidine tagged soluble form of SipS of *B. amyloliquefaciens* could be isolated in significant quantities. This protein displayed optimal activity at pH 10, which is remarkable considering the fact that the catalytic domain of SPases is located in an acidic environment at the outer surface of the membrane of living cells. Strikingly, in contrast to what has been previously reported for the soluble form of the *Escherichia coli* SPase, soluble SipS was active in the absence of added detergents. This observation can be explained by the fact that a highly hydrophobic surface domain of the *E. coli* SPase, implicated in detergent-binding, is absent from SipS.

High-level production of hexa-histidine tagged soluble forms of *Bacillus* SPases in *E. coli* was unsuccessful. Therefore, studies aiming at answering the question whether such overproduction is precluded by proteolysis were performed and are presented in Chapter 3. The results show that the degradation of soluble forms of the *Bacillus* SPase SipS is largely due to self-cleavage. First, catalytically inactive soluble forms of this SPase were not prone to degradation; in fact, these mutant proteins were produced at very high levels in *E. coli*. Second, the purified active soluble form of SipS displayed self-cleavage *in vitro*. Third, as determined by N-terminal sequencing, at least one of the sites of self-cleavage (between Ser15 and Met16 of the truncated enzyme) strongly resembles a typical SPase cleavage site. Self-cleavage at the latter position results in complete inactivation of the enzyme, as Ser15 forms the SPase catalytic dyad with Lys55. Ironically, self-cleavage between Ser15 and Met16 can not be prevented by mutagenesis of Gly13 and Ser15, which conform to the "–1, –3" rule for SPase recognition, because these residues are critical for SPase activity.

Since the *in vitro* characterisation of soluble SPase variants lacking the membrane anchor domain was severely hampered by autodegradation, intact hexa-histidine tagged variants were overproduced, purified and characterised. Chapter 4 describes the overproduction of several intact *Bacillus* SPases

to high levels. Interest
differences in the su
purified paralogous SP
be demonstrated us
substrates. Moreover,
role of the membra
substrate specificity
obtained. Finally, the m
was shown to be a m
stability of *Bacillus* SPa
autodegradation of the c

Chapter 5 documents
difference in specificity b
and SipT) and *minor* (
SPases is, at least in p
membrane anchor domai
SPases. Notably, the d
minor SPases was based
cell viability. Because th
between *major* and *mino*
reflected in sequence
phylogeny was used to p
SPases. The results were
from various bacilli. As pre
behaved as *major* or
expressed in *B. subtilis*.
modelling indicated that the
not a critical parameter for
and *minor Bacillus* SPase
substrate-binding site of the
smaller than that of other
minor SipV could be convert
without changing this site. Ir
amino-terminal residues
corresponding residues of th
was sufficient for conver
suggests that differences
minor SPases are based or
substrate cleavage site selec

Chapter 6 reports the exist
Tat pathway in *B. s*
demonstrated by studying
secretion of PhoD, a
belonging to a novel protein
known members are synth
RR-signal peptides. Unlike

to high levels. Interestingly, both similarities and differences in the substrate specificity of the purified paralogous SPases SipS and SipT could be demonstrated using different pre-protein substrates. Moreover, evidence for a possible role of the membrane anchor domain in substrate specificity or accessibility was obtained. Finally, the membrane anchor domain was shown to be a major determinant for the stability of *Bacillus* SPases *in vitro* by preventing autodegradation of the catalytic domain.

Chapter 5 documents the conclusion that the difference in specificity between the *major* (SipS and SipT) and *minor* (SipU, SipV, and SipW) SPases is, at least in part, determined by the membrane anchor domain of these two groups of SPases. Notably, the distinction in *major* and *minor* SPases was based on their importance for cell viability. Because the functional difference between *major* and *minor* SPases is not clearly reflected in sequence alignments, molecular phylogeny was used to predict *major* and *minor* SPases. The results were verified with SPases from various bacilli. As predicted, these enzymes behaved as *major* or *minor* SPases when expressed in *B. subtilis*. Strikingly, molecular modelling indicated that the active site geometry is not a critical parameter for classification of *major* and *minor Bacillus* SPases. Even though the substrate-binding site of the *minor* SPase SipV is smaller than that of other known SPases, the *minor* SipV could be converted into a *major* SPase without changing this site. Instead, replacement of amino-terminal residues of SipV with corresponding residues of the *major* SPase SipS was sufficient for conversion of SipV. This suggests that differences between *major* and *minor* SPases are based on activities other than substrate cleavage site selection.

Chapter 6 reports the existence of a functional Tat pathway in *B. subtilis*. This was demonstrated by studying the TatC-dependent secretion of PhoD, a phosphodiesterase belonging to a novel protein family of which all known members are synthesised with typical RR-signal peptides. Unlike most organisms of which the genome has been sequenced completely, the Gram-positive bacterium *B. subtilis* contains two *tatC*-like genes, denoted *tatCd* and *tatCy*. The fact that TatCd was shown to be of major importance for the secretion of PhoD, whereas TatCy is not required for this process, indicated for the first time that TatC is a specificity determinant for protein secretion *via* the Tat pathway. Based on these observations, it is hypothesised that the TatC-determined pathway specificity is based on specific interactions between TatC-like proteins and other pathway components, such as TatA, of which three paralogs are present in *B. subtilis*.

The studies described in Chapter 7 were aimed at determining the number of extracellular *B. subtilis* proteins that follow the Tat pathway by applying proteomic techniques. While most exported proteins appear to use the Sec pathway, 69 of these proteins could potentially use the Tat pathway, as their signal peptides contain RR- or KR-motifs. The results show that only the phosphodiesterase PhoD was secreted in a strictly Tat-dependent manner, whereas 13 other proteins are secreted Tat-independently, showing that their RR/KR-motifs are not recognised by the Tat machinery. In fact, the extracellular accumulation of three of these 13 proteins was shown to be SecA-dependent. The observation that the export of LipA is SecA- and not Tat-dependent is particularly remarkable, because its signal peptide conforms to the most stringent criteria for the prediction of Tat-dependent export in *E. coli*. Taken together, these observations show that the Tat pathway makes a very selective contribution to the extracellular proteome of *B. subtilis*.

Chapter 8 was aimed at answering the question whether certain SPases of *B. subtilis* have a dedicated function in the secretion of the phosphodiesterase PhoD, which is translocated *via* the Tat pathway of this organism. By using proteomic techniques it was shown that none of the SPases of *B. subtilis* is specifically required for PhoD secretion. Furthermore, the results suggest that the presence of either SipS or SipT

is sufficient for this process. Taken together, the observations presented in this Chapter support the view that the type I SPases of *B. subtilis* are functionally redundant and act in a secretion pathway-independent manner.

Notably, the secretion of only 14 of the 69 *B. subtilis* proteins with predicted RR/KR-signal peptides could be visualised by analysis of the extracellular proteome of this organism. Chapter 9 documents the use of epitope-tagging and controlled gene expression to identify additional Tat substrates of *B. subtilis*. Using these techniques, the YwbN protein was shown to be secreted Tat-dependently. This observation is consistent with the fact that YwbN is synthesised with an RR-signal peptide that conforms to the most stringent criteria for the prediction of Tat-dependency as defined for known RR-signal peptides of *E. coli*. Notably, YwbN is the first protein shown to require the TatCy component of the *B. subtilis* Tat machinery for its secretion. In contrast, the TatCd component, which is critical for the secretion of PhoD, is not required for YwbN secretion. These findings support the view that TatC is a specificity determinant for protein secretion in *B. subtilis*.

In conclusion, the studies described in this thesis provide more insights in the role of specificity determinants in protein export processes in *B. subtilis*. First, the membrane anchor domain of SPases was shown to be an important determinant for differences in activity, specificity and stability of these enzymes *in vitro* (Chapters 2-4). Moreover, the difference in specificity between the *major* (SipS and SipT) and *minor* (SipU, SipV, and SipW) SPases was shown to be, at least in part, determined by this membrane anchor domain (Chapter 5). Second, it was shown that signal peptides determine pathway choice, but that the Tat pathway is very selective in the transport of proteins containing potential RR/KR-motifs in their signal peptides (Chapters 6 and 7). Third, by studying the secretion of two different Tat substrates, the TatC component of *B. subtilis* was shown to be a specificity determinant for protein secretion (Chapters 6 and 9). Finally, no Tat-dedicated *B. subtilis* SPase could be identified, as the presence of either the *major* SipS or SipT protein was shown to be sufficient for the processing of PhoD (Chapter 8).